AMENDMENT TO THE SPECIFICATION

Please replace the paragraph beginning on page 12, line 22 to page 13, line 10 with the following amended paragraph:

It should be noted that the entire system 100, or part of system 100 can be implemented in the environment illustrated in FIG. 1. Feature extraction module 106 and trainer module 108 can be either hardware modules in computer 20 or software modules stored in any of the information storage devices disclosed in FIG. 1 and accessible by CPU 21 or another suitable processor. In addition, lexicon storage module 109, acoustic models 111, and language models 110 are also preferably stored in any of the suitable memories devices shown in FIG. 1. Further, search engine 107 can be implemented in CPU 21, which can include one or more processors or can be performed by a dedicated speech recognition processor employed by personal computer 20. In addition, output device ~~112~~115 and I/O device ~~113~~116 can include any of the I/O devices shown in FIG. 1, such as keyboard 40, pointing device ~~43~~42, monitor 47, a printer or any of the memory devices shown in FIG. 1, for example.

Please replace the paragraph beginning on page 13, line 18 to page 14, line 4 with the following amended paragraph:

Feature extraction module 106 divides the digital signals into frames, each of which includes a plurality of digital samples. In one embodiment, each frame is approximately 10 milliseconds in duration. The frames are then encoded into a feature vector reflecting the spectral characteristics for a plurality of frequencies bands. In the case of ~~discreet~~discrete and semi-continuous hidden Markov modeling, feature extraction model 106 also encodes the feature vectors into one or more code words using vector quantization techniques and a codebook derived

from training data. Thus, feature extraction module 106 provides, at its output, the feature vectors (or codewords) for each spoken utterance. Feature extraction module 106 preferably provides the feature vectors at a rate of one feature vector approximately every 10 milliseconds, for example.

Please replace the paragraph beginning on page 14, line 14 to page 15, line 8 with the following amended paragraph:

The stream of feature vectors produced by feature extraction module 106 is provided to speech recognizer 107, which identifies a most likely sequence of speech units, such as words or phonemes, based on the stream of feature vectors, one or more acoustic models in repository 111, one or more of language models in repository 110, and lexicon ~~105~~109. Caller identification module 112 identifies the caller as a new caller or one of any previously identified callers, by applying the feature vectors of the voice input to generic and caller-specific models of the speech units identified by speech recognizer 107, which are stored in repository 111. In one embodiment, caller identification module 112 also uses generic and caller-specific language models, stored in repository 110, to assist in the identification. Module 112 outputs the caller identity and/or text of the most likely sequence of uttered words to call router 113 or stores these results in one of the memory devices shown in FIG. 1, for example. The results can also be output to a user or operator through I/O device 115. Call router 113 can then screen the call or route the call to one or more selected destinations based on the identity of the caller and/or the content of the call.